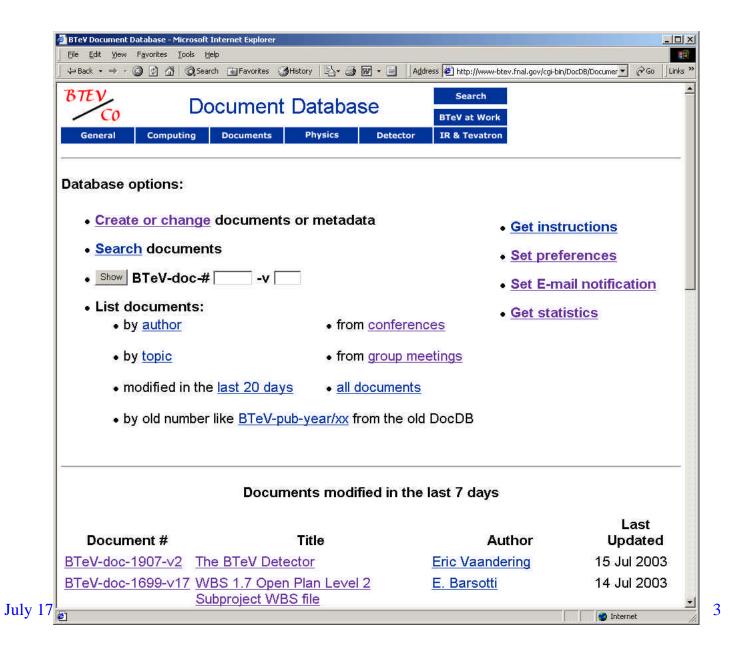# Document Management at BTeV
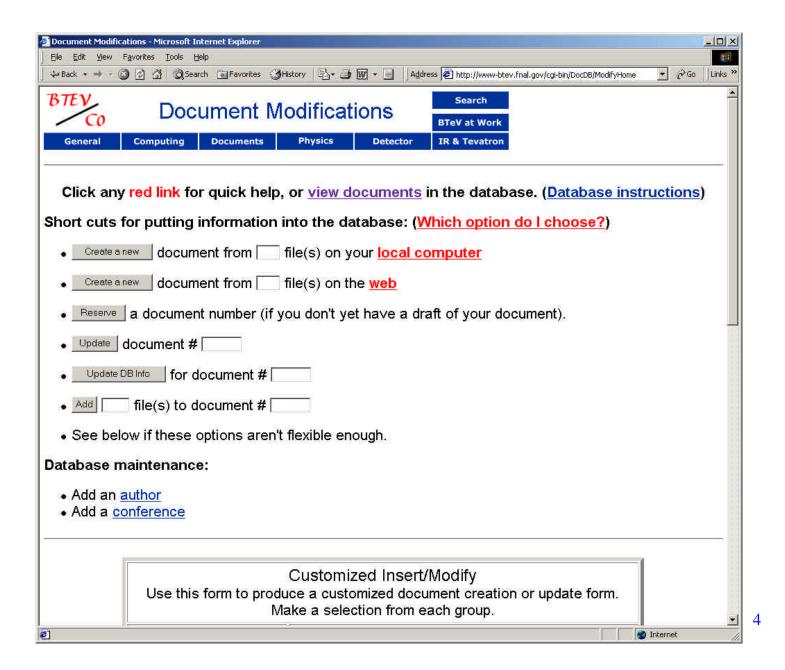
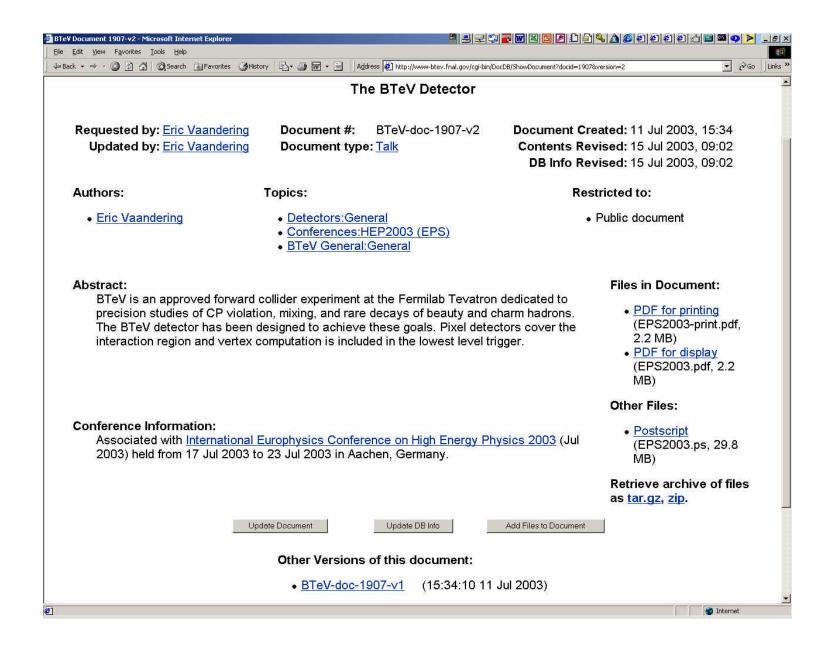## Rob Kutschke

## July 17, 2003

Public access to the BTeV document database:
http://www-btev.fnal.gov/cgi-bin/public/DocDB/DocumentDatabase

# What is a Document?

- User view: Everything is a document.
  - Text, figures, digital photos, sound, movies, CAD drawings, blueprints, html, scanned scribbling on a napkin, a note that Lory keeps our copy of the book …
  - Conference talks, publications, group meeting talks, posters, EOI, Proposal, TDR, internal notes, email from lab director, vendor quotes, living documents …
- Technical view: Meta-data + files.
  - "Unlimited" number and types of files per document.
  - "Unlimited" number of versions per document.
    - All versions retained.

# BTeV Document Database — Microsoft Internet Explorer

File   Edit   View   Favorites   Tools   Help

Back   Search   Favorites   History   Address   http://www-btev.fnal.gov/cgi-bin/DocDB/Documer   Go   Links »

**BTeV Co**

## Document Database

**Search**
**BTeV at Work**

| General | Computing | Documents | Physics | Detector | IR & Tevatron |

## Database options:

- **Create or change** documents or metadata

- **Search** documents

- [Show] **BTeV-doc-#** [      ] **-v** [      ]

- **List documents:**
  - by author
  - by topic
  - modified in the last 20 days
  - by old number like BTeV-pub-year/xx from the old DocDB

  - from conferences
  - from group meetings
  - all documents

- **Get instructions**

- **Set preferences**

- **Set E-mail notification**

- **Get statistics**

---

### Documents modified in the last 7 days

| Document # | Title | Author | Last Updated |
|---|---|---|---|
| BTeV-doc-1907-v2 | The BTeV Detector | Eric Vaandering | 15 Jul 2003 |
| BTeV-doc-1699-v17 | WBS 1.7 Open Plan Level 2 Subproject WBS file | E. Barsotti | 14 Jul 2003 |

Internet

# The BTeV Detector

**Requested by:** Eric Vaandering    **Document #:**   BTeV-doc-1907-v2    **Document Created:** 11 Jul 2003, 15:34

**Updated by:** Eric Vaandering    **Document type:** Talk    **Contents Revised:** 15 Jul 2003, 09:02

**DB Info Revised:** 15 Jul 2003, 09:02

**Authors:**

- Eric Vaandering

**Topics:**

- Detectors:General
- Conferences:HEP2003 (EPS)
- BTeV General:General

**Restricted to:**

- Public document

**Abstract:**

BTeV is an approved forward collider experiment at the Fermilab Tevatron dedicated to precision studies of CP violation, mixing, and rare decays of beauty and charm hadrons. The BTeV detector has been designed to achieve these goals. Pixel detectors cover the interaction region and vertex computation is included in the lowest level trigger.

**Files in Document:**

- PDF for printing (EPS2003-print.pdf, 2.2 MB)
- PDF for display (EPS2003.pdf, 2.2 MB)

**Other Files:**

- Postscript (EPS2003.ps, 29.8 MB)

**Conference Information:**

Associated with International Europhysics Conference on High Energy Physics 2003 (Jul 2003) held from 17 Jul 2003 to 23 Jul 2003 in Aachen, Germany.

**Retrieve archive of files as** tar.gz, zip.

Update Document    Update DB Info    Add Files to Document

**Other Versions of this document:**

- BTeV-doc-1907-v1    (15:34:10 11 Jul 2003)

# Meta-Data for 1 Version of 1 Doc

- Title
- Requestor
- Author(s)
- Dates of creation and modifications.
- Document type ( note, talk, poster … )
- Abstract

- Security info.
- Topics and keywords
- Short comment about each file in the doc.
- If applicable:
    - Publication info.
    - Link to conference info.

# The Underlying Technology

- Meta-data: MySQL database.
- Documents are files in a unix filesystem.
  - Actual files are stored, not urls.  No link rot.
  - One subdirectory per version of each doc.
- Web interface for viewing, adding, modifying and searching documents (Perl/cgi).
  - Has detailed online help.
  - Interface is very, very well done.
- Web interface for administrator functions.
  - Admins do not need to know SQL etc.

# Security: Nested Groups

| Executive | | | Public | Admin |
|---|---|---|---|---|
| BTeV (default) | | | | |
| RTES | Review (read only) | | | |

- .htaccess (per doc version).
- Group password.
- Parent group may read files of child group.
- One doc may have both RTES and Review authorizations.

- Review account enabled only during reviews.
- Separate authorization for read and for add/modify.
- System is extensible – can add new boxes at any level.

# Does not Replace Web Pages

- This does not replace handcrafted web pages.
  - When a list of relevant documents only makes sense if supplied with context information.
  - When you want to highlight a particular set of documents.
    - BTeV: "Proposal, TDR and related documents".
- Can link from a web page to:
  - The meta-data display for a document.
  - Any file in a document with version given as:
    - the highest version of a document.
    - a specified version of a document.
    - the version which was current at a specified date in the past.
  - The search engine.

# Social Engineering

- This is as important as the software engineering.
- People need to use this tool in a way that others will understand.
- Maintain topics and keywords, educate people in their use, and get people to use them.
- File formats: if someone uses a format which not all can read, they should also provide a .pdf copy.
- Are a conference talk and its associated proceedings one document or two?

# Structure vs Flexibility

- Must get this tradeoff right.
- Topics:
  - Administratively defined.  Every document must have one or more topics from the predefined list.
- Keywords:
  - Free form.  No administrative control
  - Subgroups develop policies for their own internal use. Conformance controlled by peer pressure.
- BTeV: early on there were lots of revisions in the topic list.  Working well now.

# User Experience

- BTeV had a killer app:
  - Great utility for meetings with multi-way video conferencing.
  - Quickly accepted by the collaboration.
    - Learning curve was fast.
  - Reviewers liked it.

- Beams division played with a fancier security scheme, gave that up and quickly adopted it (almost?) as is.

# Support Requirements

- Technical support: O(1 FTE day) to setup +  O(a few FTE hours/month) to maintain ( assuming an existing server ).
  - Configure the DocDB:
    – Create list of  authors/requestors ( hours )
    – Need input from throughout the division, both low and high level staff. (5-10 people for ~1 hour, ~5 times? )
      - Any changes to the meta data?
      - List of topics and guidelines for keywords.
      - Define security groups.
    – Ongoing component: (5-10 people for ~1 hour, a few times/year )?

- User education at startup  ( a few FTE weeks )
    – Should decay away with a half life of weeks.

- Adding new features will take more effort (???).

- What if we port to a non-linux server: ( Used to run on OSF and conversion to linux was easy).

# In the works

- Better support for meeting agendae.
- For other stuff, check with EricVaandering ( ewv@fnal.gov ).

# Wish List

- Full text search ( now only meta-data is searchable ).
    - Fully general solution is hard ( how do you search postscript?).
    - Possible to use htdig as a plugin?
- Automagic conversion to PDF from proprietary formats so that users without the proprietary software can view the document.
    - Adobe offers a service to do this.

# Questions and Recommendations

- Does the security model meets the needs of the division?
  - Most variations I can think of would require more resources to keep it up to date.
- I strongly recommend this as the way to maintain CD documents for CD people and their non-CD co-workers.
- I do not recommend it, as is, as the sole way to present any CD document to non-CD people.
  - My guess is that we still need some hand maintained web pages for documents accessed either frequently or primarily by new users.
    - Such web pages should link into documents which live in the DocDB.
    - The web page is just a customized view of some subset of the docs.
    - This gives us the best of both worlds.

# What about the Wiki

- A public scribble pad.
  - Old versions are kept but access is a little clumsy.
  - No meta data, only the living document.
  - Great for free form discussion of developing ideas
    - Not everyone can work this way.
    - Not as good for work which requires externally imposed structure to maintain order.

- This is not a document management tool.
  - It solves a totally different problem.

# Backup Slides

# Plugs

- Main work done by Eric Vaandering (Vanderbilt ) and Lynn Garren.

- Many ideas but no code taken from NuMI Notes (Liz Buckley-Geer ).

- I think that this is well designed, well executed and it has made a big improvement to the way we work at BTeV.

# Miscellaneous Comments

- Intro page has the last n days worth of stuff.

- Email alerts

- Some interfaces can be customized using cookies.

- Reviewers like it.

- Searches return info only about docs you may see.

# Security

- Web server username/password to grant access.
  - "Group password" style.
- Granularity: per version of document.
- BTeV's original security heirarchy:
  - Access to files:
    - Executive – may view any document.
    - BTeV　　　– may view any non-Executive document.
    - Public　　– may view only public documents.
  - DocDBAdmin – may perform admin functions.

# Security …

- General model: nested subgroups, each more restricted.
- BTeV used this model to define two new security levels as subgroups of the BTeV group.
  - RTES
    - Document may be viewed/modified using any of the RTES, BTeV or Executive passwords.
  - Review
    - Document may be viewed using any of the Review, BTeV or Executive passwords.
    - Reviewers may not add/modify documents.
    - The review password is given to authorized reviewers for the duration of a review and it is inactivated after the review has concluded.

# Support – Startup (1)

- Hardware: depending on the load, this can run on an existing server. ( zero to small additional work )
- Software: install web server S/W, MySQL, perl. (small)
- Install DocDB software. (hours)
- Configure DocDB:
  - Create list of legal authors/requestors ( hours )
  - Decide on details. Need input from throughout the division, both low and high level staff. ( 5-10 people for 1 hour, 5 times? )
    - Any changes to the meta data?
    - List of topics and guidelines for keywords.
    - Define security group heirarchy.

# Support – Startup (2)

- Staff training ( a few FTE weeks plus an hour per general user? )
  - Train administrators.
  - Write two page intro document for general users.
  - May need to spin the online help differently.
  - BTeV experience may not be relevant.
- Entry of legacy documents ( huge effort if there are a lot of legacy documents but it can be done by secretarial staff, as was done for BTeV ).

# Support - Ongoing

- Maintain hardware and underlying software.  Backups.
  - Zero to  small additional work ?
- Install updates of DocDB software  ( small, unless you want a lot of new features )
- Maintain author and topic lists ( 5-30 secretary minutes/day ?? )
- Bug and user-screwup fixing ( small at BTeV ).
- A more sophisticated security scheme might well have greater support requirements. (???).
- Handholding/ongoing training (at BTeV this had a half-life of weeks – everyone learned from their friend)